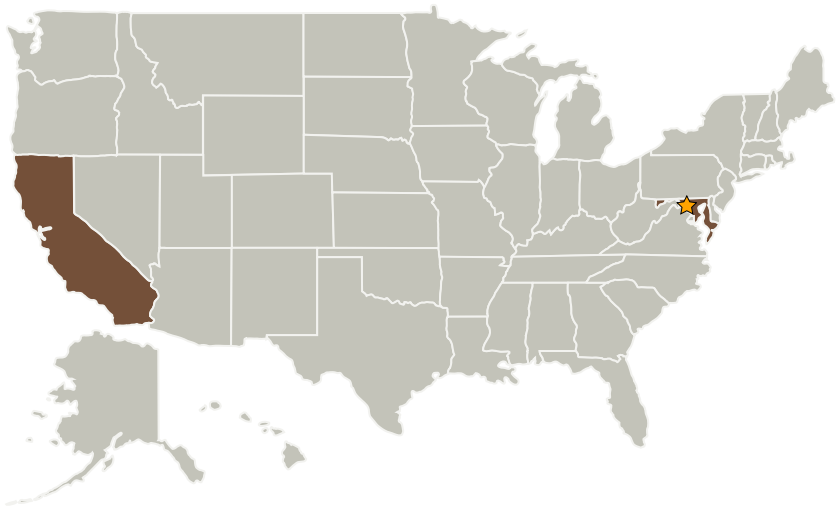
# Computing Infrastructure and Remote, Parallel Data Mining Engine for Virtual Observatories, Phase I

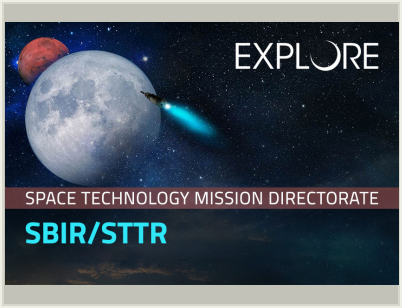
Completed Technology Project (2005 - 2005)

## Project Introduction

We propose to develop a state-of-the-art data mining engine that extends the functionality of Virtual Observatories (VO) from data portal to science analysis resource. Our solution consists of two integrated products, IDDat and RemoteMiner: (1) IDDat is an advanced grid-based computing infrastructure which acts as an add-on to VOs and supports processing and remote data analysis of widely distributed data in space sciences. IDDat middleware design is such as to reduce undue network traffic on the VO. (2) RemoteMiner is a novel data mining engine that connects to the VO via the IDDat. It supports multi-users, has autonomous operation for automated systematic identification while enabling the advanced users to do their own mining and can be used by data centers for pre-mining. These innovations will significantly enhance the science return from NASA missions by providing data centers and individual researchers alike an unprecedented capability to mine vast quantities of data. Phase I is aimed at complete definition of the design of the product and a demonstration of a prototype of the proposed major innovations. Phase II work will encompass the building of a full commercial product with associated production quality technical and user documentation.

## Primary U.S. Work Locations and Key Partners






Computing Infrastructure and Remote, Parallel Data Mining Engine for Virtual Observatories, Phase I

## Table of Contents



## Organizational Responsibility

**Responsible Mission Directorate:**

Space Technology Mission Directorate (STMD)

**Lead Center / Facility:**

Goddard Space Flight Center (GSFC)

**Responsible Program:**

Small Business Innovation Research/Small Business Tech Transfer

# Computing Infrastructure and Remote, Parallel Data Mining Engine for Virtual Observatories, Phase I

Completed Technology Project (2005 - 2005)

| Organizations Performing Work | Role | Type | Location |
|---|---|---|---|
| ★Goddard Space Flight Center(GSFC) | Lead Organization | NASA Center | Greenbelt, Maryland |
| Scibernet, Inc. | Supporting Organization | Industry | San Diego, California |

| Primary U.S. Work Locations | |
|---|---|
| California | Maryland |

## Project Management

**Program Director:**

Jason L Kessler

**Program Manager:**

Carlos Torrez

**Principal Investigator:**

Homa Karimabadi

## Technology Areas

**Primary:**

- TX11 Software, Modeling, Simulation, and Information Processing
  - TX11.6 Ground Computing
    - TX11.6.7 High Performance Data Analytics Platform